NRL Report 6217
NAREC Reference 31

Tests on the Multiplicative Congruence Method of Generating Pseudo-Random Numbers on the NAREC Computer

DORIS J. ELLIS AND P. C. RYAN

Applied Mathematics Staff
Office of Director of Research

April 2, 1965

U.S. NAVAL RESEARCH LABORATORY
Washington, D.C.

CONTENTS

ABSTRACT

This report describes the statistical tests and their results that were made on pseudo-random numbers generated by the "multiplicative congruential method" on the NAREC computer. The first two million numbers of a particular generated sequence were tested using the frequency test, poker test, serial test, runs above and below the mean test, runs up and down test, and serial correlation test. The purpose of making such tests was to determine if the generated numbers were, in fact, random in their behavior. It was found that these numbers can be considered random for most common applications.

PROBLEM STATUS

This is a final report on one phase of the problem; work on this problem is continuing.

AUTHORIZATION

NRL Problem F04-05
Project RR 009-03-45-5802

Manuscript submitted November 16, 1964.

TESTS ON THE MULTIPLICATIVE CONGRUENCE METHOD OF GENERATING PSEUDO-RANDOM NUMBERS ON THE NAREC COMPUTER

INTRODUCTION

The following report presents the results of several tests of randomness applied to pseudo-random numbers which were generated by the "multiplicative congruential method" (1).

As implemented on the NAREC, this method produces a sequence $x_0, x_1, \ldots, x_n$ of positive integers satisfying the recursion formula

$$x_i \equiv 5^{17} x_{i-1} (\text{mod } 2^{44})$$

where this notation means that x_i is the remainder when $5^{17}x_{i-1}$ is divided by 2^{44}. (Thus, $0 \leq x_i < 2^{44}$.) Pseudo-random numbers (between 0 and 1) are then obtained by dividing each x_i by 2^{44}.

The recursive computations can be done easily by low-order multiplication on the NAREC. When two binary integers, a and b (44 bits each), are multiplied, the result is

$$ab = c_1 2^{87} + c_2 2^{86} + \ldots + c_{87} 2^1 + c_{88} 2^0$$

where $c_1, \ldots, c_{88}$ are the binary digits 0 or 1. On dividing such a product by 2^{44}, we obtain

$$\frac{ab}{2^{44}} = \frac{c_1 2^{87} + \ldots + c_{44} 2^{44}}{2^{44}} + \frac{c_{45} 2^{43} + \ldots + c_{88} 2^0}{2^{44}}$$

$$= \frac{(c_1 2^{43} + \ldots + c_{44} 2^0) \cancel{2^{44}}}{\cancel{2^{44}}} + \frac{c_{45} 2^{43} + \ldots + c_{88} 2^0}{2^{44}}$$

$$= k + \frac{c_{45} 2^{43} + \ldots + c_{88} 2^0}{2^{44}}$$

where k is an integer, and $0 \leq (c_{45} 2^{43} + \ldots + c_{88} 2^0) < 2^{44}$. That is, the integer $c_{45} 2^{43} + \ldots + c_{88} 2^0$ is the remainder when the product, ab, is divided by 2^{44}, or

$$c_{45} 2^{43} + \ldots + c_{88} 2^0 \equiv ab (\text{mod } 2^{44}) .$$

But, in the NAREC the digits $c_{45}, \ldots, c_{88}$ are simply the contents of the U register.

A sequence generated in this fashion is completely determined by its starting value, x_0, which must be an odd integer. There are two essentially different sequences possible:

(a) If x_0 is one of the integers 1, 5, 9, 13, . . . , then so is every x_i, and each sequence $\{x_i\}$ consists precisely of all such integers less than 2^{44} and arranged in some fixed cyclic order. Two type (a) sequences with unequal starting values differ only in that each is a cyclic permutation of the other.

(b) If x_0 is one of the integers 3, 7, 11, 15, . . . , then again every x_i is also, and the sequences $\{x_i\}$ consist of all such integers less than 2^{44}, in a fixed cyclic order. Again, any two sequences of this type with unequal starting values differ only in that each is a cyclic permutation of the other.

In either the type (a) or the type (b) sequences there are $2^{44}/4 = 2^{42} \approx 4 \times 10^{12}$ distinct entries before the first repetition, after which the sequences begin to cycle with period 2^{42}. (This is not the way true random numbers behave, of course. However, the problem of periodicity does not arise in the statistical tests performed here, since all samples are based on the first two million (or less) numbers of the sequence beginning with $x_0 = 1$.)

The pseudo-random number between 0 and 1, which corresponds to the integer x_i, will be denoted by R_i. Thus, $R_i = x_i/2^{44}$. (For example, $R_0 = 1/2^{44}$ and $R_1 = 5^{17}/2^{44}$.) Each R_i has an exact representation in the hexidecimal number system: namely, a hexidecimal point followed by 11 hexidecimal digits. It is convenient in what follows to think of the generated numbers as represented in this form.

Samples of k numbers to be tested are taken from the generated sequence $\{R_i\}$ according to one of the following schemes:

$$R_1, R_2, \ldots, R_k$$

$$R_1, R_3, \ldots, R_{2k-3}, R_{2k-1}$$

$$R_2, R_4, \ldots, R_{2k-2}, R_{2k}.$$

In accordance with terminology currently in use, the above sampling procedures will be referred to as "lag 1," "lag 2 beginning with R_1," and "lag 2 beginning with R_2," respectively. (The notation $\{N_j\} = N_1, N_2, \ldots, N_k$ will be used throughout to denote any of the above subsequences of the generated sequence $\{R_i\}$.)

The statistical tests used in this report are among those most commonly used (see Contents). All statistical tests for randomness are based upon probability statements that can be made concerning randomly selected objects, and a sample of random numbers would be expected to conform approximately to such probability statements.

One of the most widely used and most efficient statistical tests to determine the degree of compatibility between observed and expected frequencies (goodness of fit) is the chi-square (χ^2) test. A brief description of this test is as follows:

Suppose each of n pieces of data is assigned to one of k mutually exclusive classes. Let f_i be the observed frequency of the ith class. If the probability of any single piece of data falling into the ith class is a known constant p_i, where

$$\sum_{i=1}^{k} p_i = 1,$$

then the expected frequency of the ith class is given by $e_i = np_i$. In this situation the statistic

$$X^2 = \sum_{i=1}^{k} \frac{(f_i - e_i)^2}{e_i} = \sum_{i=1}^{k} \frac{(f_i - np_i)^2}{np_i} \qquad (1)$$

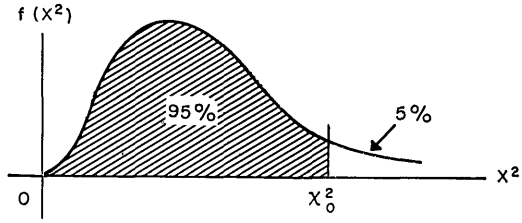
has (as $n \to \infty$) the χ^2 distribution with $k - 1$ degrees of freedom* (2a). When the p_i are unknown, but can be estimated from the sample by the method of (multinomial) maximum likelihood, X^2 can be shown to possess the χ^2 distribution with $k - s - 1$ degrees of freedom, where s is the number of estimated independent parameters (2b).

In each goodness-of-fit test used here, a hypothesis is formulated concerning the p_i. This hypothesis may be very specific, assigning exact values to the probabilities (e.g., $p_i = 1/k$, where $i = 1, \ldots, k$), or it may only impose a mild restriction on the values which some of the p_i's may assume (e.g., $p_2 = 2p_1$, where $p_3, \ldots, p_k$ are unspecified). In either case, assuming that the hypothesis is true, the X^2 statistic has (asymptotically) a particular χ^2 distribution (depending on the number of degrees of freedom involved). Thus a particular value, χ_0^2, may be determined such that

$$P\{X^2 < \chi_0^2\} = .95$$

The computed value of X^2 will be less than χ_0^2 in 19 out of 20 such tests if the hypothesis is actually true. The situation is as depicted in Fig. 1. If the computed value of X^2 falls to the left of χ_0^2 in the figure, we agree to accept the hypothesis; if it falls to the right (i.e., in the "critical region") the original hypothesis is rejected.

Fig. 1 - The χ^2 distribution



In this report the upper-tail (one-sided) test is used, rather than the two-tail test. The problem of choosing between the upper-tail and the two-tail test is discussed by Kendall and Stuart (2c).

The χ_0^2 value is determined from a table of the "percentage points of the χ^2 distribution," using degrees of freedom versus percentage level (.95 in this report) as the point of reference. For degrees of freedom greater than 30 or 100 (depending on length of table), it is efficient to use the approximation

$$\chi^2_{d,p} = d\left[1 - \frac{2}{9d} + Z_p\left(\frac{2}{9d}\right)^{1/2}\right]^3$$

where d is the degrees of freedom and Z_p is the (100 p)th percentile of the standard unit normal distribution. (In this report $p = .95$ and $Z_p = 1.6449$.)

*Degrees of freedom can be thought of as the number of classes whose frequencies can be assigned without restriction, or the number of independent frequency pairs being compared.

The next section of this report deals with the kinds of statistical tests which were applied. Description of specific tests and their results will appear in the Test Results section.

TESTS FOR LOCAL RANDOMNESS

The Frequency Test

The frequency test requires that the distinct elements of a set occur approximately an equal number of times. The expected number of occurrences of a specific element is n/(number of distinct elements), where n is the total number of occurrences of all elements. Degrees of freedom equals the number of distinct elements in the set minus one.

The Poker Test

The poker test requires that if the digits of a series are taken in blocks of five digits, there will be a certain expectation of the number of blocks in which the following combinations occur:

1. Bust - all digits different, denoted by abcde
2. One pair, denoted by aabcd
3. Two pair, denoted by aabbc
4. Three of a kind, denoted by aaabc
5. Full house, denoted by aaabb
6. Four of a kind, denoted by aaaab
7. Five of a kind, denoted by aaaaa.

The poker test checks mutual independence of the various digit positions. If random selection is assumed, the following probabilities hold for the above combinations: n = number of different digits (10 in a decimal system, 16 in a hexadecimal system, etc.)

1. $P(\text{abcde}) = n(n-1)(n-2)(n-3)(n-4)/n^5$
2. $P(\text{aabcd}) = n(n-1)(n-2)(n-3)\binom{5}{2}/n^5$
3. $P(\text{aabbc}) = \binom{n}{2}(n-2)\binom{5}{2}\binom{3}{2}/n^5$
4. $P(\text{aaabc}) = n(n-1)(n-2)\binom{5}{3}/n^5$
5. $P(\text{aaabb}) = n(n-1)\binom{5}{2}/n^5$
6. $P(\text{aaaab}) = n(n-1)\binom{5}{4}/n^5$
7. $P(\text{aaaaa}) = n/n^5$.

The number of degrees of freedom is the number of combinations minus one.

The Serial Test

The serial test requires that a generated number of one magnitude shall not tend to be followed by or associated with a number of any other specific magnitude. If the interval 0 to 1 is divided into I equal intervals, a contingency (bivariate) table may be constructed showing the distribution of interval pairs for adjacently occurring numbers in a given sequence of random numbers. The table is arranged in rows according to the first interval of each pair, and in columns according to the second. Approximately equal frequencies should be obtained in all cells.

To assist in insuring independence of interval pairs, the totality of pairs of adjacent numbers is split into two samples in the following manner:

Given a sequence of numbers $N_1, N_2, \ldots, N_{k-1}, N_k$ (k even), the first test uses the following pairs:

$$(N_1, N_2);\ (N_3, N_4);\ \ldots;\ (N_{k-1}, N_k)\ .$$

The second test then uses the remaining pairs, as follows:

$$(N_2, N_3);\ (N_4, N_5);\ \ldots;\ (N_{k-2}, N_{k-1});\ (N_k, N_1)\ .$$

For each of the above tests, the sample size is $k/2$.

In each test the contingency table yields a two-way classification of the frequencies. As in Eq. (1) the statistic

$$X^2 = \sum_{i=1}^{I}\sum_{j=1}^{I}\frac{(f_{ij}-e_{ij})^2}{e_{ij}} = \sum_{i=1}^{I}\sum_{j=1}^{I}\frac{(f_{ij}-np_{ij})^2}{np_{ij}}$$

has the χ^2 distribution with I^2-1 or I^2-s-1 degrees of freedom, where s is the number of independent parameters estimated, I is the number of rows or columns in the contingency table, and n is the sample size. The number s is determined by the hypotheses tested. Four χ^2 tests are used on each sample of interval pairs, with the following meanings:

$\chi^2(1)$ - This test examines the independence of sequentially occurring intervals. An acceptable $X^2(1)$ value is an indication that the occurrence of an interval does not alter the probabilities of interval occurrence for the next random number generated. The hypothesis to be tested is as follows:

$$H_0:\ p_{ij} = p_{i.}p_{.j} \begin{cases} i = 1, \ldots, I \\ j = 1, \ldots, I \end{cases}$$

where

$$p_{i.} = \sum_{j=1}^{I} p_{ij} \quad \text{and} \quad p_{.j} = \sum_{i=1}^{I} p_{ij}\ .$$

Here the $p_{i.}$ and $p_{.j}$ are unknown and must be estimated. Their maximum likelihood estimates are $\hat{p}_{i.} = r_i/n$ and $\hat{p}_{.j} = c_j/n$, where r_i is the row total, c_j is the column total, and n is the sample size. Hence, under H_0, the maximum likelihood estimates of the p_{ij} are $\hat{p}_{ij} = \hat{p}_{i.}\hat{p}_{.j} = r_ic_j/n^2$, and similarly $\hat{e}_{ij} = n\hat{p}_{ij} = r_ic_j/n$. Since

$$\sum_{i=1}^{I} \sum_{j=1}^{I} p_{ij} = 1$$

it follows that

$$\sum_{i=1}^{I} p_{i.} = 1 \quad \text{and} \quad \sum_{j=1}^{I} p_{.j} = 1.$$

Therefore there are $s = (I-1) + (I-1) = 2I - 2$ independent parameters to be estimated. Thus, the number of degrees of freedom (d.f.) is given by $I^2 - (2I-2) - 1 = I^2 - 2I + 1 = (I-1)^2$. (See Ref. 3 or Ref. 4a.)

$\chi^2(2)$ - A random number in one interval is equally likely to be followed by a random number in any other interval. The hypothesis is

$$H_0: p_{ij} = p_{i.} \frac{1}{I}.$$

Here the $p_{i.}$ must be estimated. Since $\hat{p}_{i.} = r_i/n$, then $\hat{p}_{ij} = r_i/nI$ and $\hat{e}_{ij} = r_i/I$. Also, $s = I - 1$ and thus $\text{d.f.} = I^2 - (I-1) - 1 = I^2 - I$.

$\chi^2(3)$ - A random number in one interval is equally likely to be preceded by a random number in any other interval. The hypothesis is

$$H_0: p_{ij} = \frac{1}{I} p_{.j}.$$

Here the $p_{.j}$ must be estimated. Since $\hat{p}_{.j} = c_j/n$, then $\hat{p}_{ij} = c_j/nI$ and $\hat{e}_{ij} = c_j/I$. Also, $s = I - 1$ and thus $\text{d.f.} = I^2 - (I-1) = I^2 - I$.

$\chi^2(4)$ - All pairs in the contingency table are equally likely. The hypothesis is

$$H_0: p_{ij} = \frac{1}{I}\frac{1}{I} = \frac{1}{I^2}.$$

Here $e_{ij} = n/I^2$. Since there are no parameters to be estimated, $\text{d.f.} = I^2 - 1$.

The Runs Above and Below the Mean Test

If k successive numbers are all greater than or all less than the mean and both the preceding and following numbers are respectively less than or greater than the mean, this is tallied as a run of length k. In a given series of n numbers there is a specific expectation, n_k, of the number of runs of length k.

$$n_k = \frac{n}{2^{k+1}}.$$

In this particular test, the χ^2 test cannot be used because the probabilities of occurrence, for different length runs, are not independent. If the actual results and the expected results are "fairly close," the random numbers are said to satisfy this test.

The Runs Up and Down Test

If k successive differences of successive numbers have the same sign (all k differences are positive or all k differences are negative), and both the preceding and following

differences are of the opposite sign, this is tallied as a run of length k, which has a certain expectation of occurrence in a given series of numbers. This expectation has been derived by Kendall (5) as follows:

The probability of run length k in a series of n numbers is

$$6(k^2 + 3k + 1)(n - k - 2)/(k + 3)!\ (2n - 7)\ .$$

The expected total number of runs is $(2/3)(n - 2)$. Here, as in the runs above and below the mean test, the χ^2 test cannot be used.

The Serial Correlation Test (See Ref. 4b)

If a set of observations is ordered with respect to time and if time is irrelevant to the variable being discussed, no correlation should exist between successive pairs of values in the sequence. Therefore, successive pairs of random numbers in the random number sequence should be independent (uncorrelated).

For k generated random numbers N_j, there are k pairs entering into the correlation, and the variates N_j and N_{j+1} differ only by order (i.e., the first refers to the numbers $N_1, N_2, \ldots, N_k$ and the second to the numbers $N_2, N_3, \ldots, N_k, N_1$).

No correlation exists if the serial correlation coefficient r is zero.

$$r = \frac{\sum_{j=1}^{k} N_j N_{j+1} - k\bar{N}^2}{k s_N^2}$$

where $\bar{N}$ and s_N^2 are the mean and the variance of the k generated random numbers. In the interpretation of a computed value of r, it is wise to keep in mind the following two examples:

(a) In the sequence $[.01, .02, \ldots, .99]$, $r = +1$.

(b) In the sequence $[.25, .75, .25, .75, \ldots, .25, .75]$, $r = -1$.

Ideally, the serial correlation coefficient should be computed for all $k!$ possible permutations of the random number sequence, but this is prohibitive for k at all large. If it is assumed that all permutations are equally probable, the distribution of the serial correlation coefficient can be approximated.

The only quantity in r that is affected by the $k!$ permutations of the sequence is the sum

$$\sum_{j=1}^{k} N_j N_{j+1}$$

and therefore it suffices to consider the statistic

$$Q = \sum_{j=1}^{k} N_j N_{j+1}\ .$$

Under the stated assumption and the hypothesis of zero correlation, the random variable Q possesses an approximate normal distribution (for large k) with mean

$$E(Q) = \frac{S_1^2 - S_2}{k - 1}$$

and variance

$$\sigma_Q^2 = \frac{S_2^2 - S_4}{k - 1} + \frac{S_1^4 - 4S_1^2 S_2 + 4S_1 S_3 + S_2^2 - 2S_4}{(k-1)(k-2)} - E^2(Q)$$

where $S_p = N_1^p + N_2^p + \ldots + N_k^p$. Then the expression

$$t = \frac{Q - E(Q)}{\sigma_Q},$$

transforms the random variable Q, which possesses a normal distribution with mean $E(Q)$ and variance σ_Q^2, to a random variable t, which possesses the normal distribution with mean zero and variance one. (This transformation is made because existing tables are for a normal curve with mean equal to zero and variance equal to one.)

Testing the hypothesis that $t = 0$ is then equivalent to testing for zero serial correlation in the original sequence of generated numbers. The critical region consists of the two symmetric tails which, taken together, measure 5 percent of the total area under the normal curve which has mean zero and variance one. Thus, the random number sequence satisfies the hypothesis of zero correlation if t is between $t_0 = -1.96$ and $t_1 = 1.96$, the normal deviates which bound 2.5 percent of the area at each end of the tabulated normal curve. These normal deviates are taken from a table of "normal areas and ordinates."

TEST RESULTS

Some of the following tests (those using 36 intervals between 0 and 1) were made for a particular problem that utilizes these random numbers. Unless otherwise stated, all tests use lag 1 (every number) beginning with R_1.

Frequency Tests

Test 1 - The first five digit positions of each random number were checked for frequency of occurrence of each of the 16 hexadecimal digits (0 through f). This was done for four sequences of random numbers: 10,000, 50,000, 100,000, 500,000 random numbers. Using the χ^2 test with 15 degrees of freedom, the 95 percent critical value is $\chi_0^2 = 24.99$. The X^2 values for Test 1 are:

Sample Size	X^2 for First Five Digit Positions				
	1	2	3	4	5
10,000	22.67	9.89	16.06	19.66	17.24
50,000	16.15	12.50	14.08	<u>26.81</u>	17.40
100,000	15.43	12.65	11.77	12.26	23.32
500,000	18.43	<u>27.35</u>	11.43	8.01	20.69

The values of X^2 that are underlined are larger than χ_0^2.

Test 2 - The interval 0 to 1 was divided into 36 equal intervals. The distribution of the random numbers in these 36 intervals was checked for each of four sequences of random numbers: 10,008, 50,004, 100,008, and 500,004 random numbers. Using the χ^2 test with 35 degrees of freedom, the 95 percent critical value is $\chi_0^2 = 49.80$. The X^2 results are:

Sample Size	X^2
10,008	44.12
50,004	34.68
100,008	28.33
500,004	35.31

all of which are less than χ_0^2.

Test 3 - Again the interval 0 to 1 was divided into 36 equal intervals. The distribution in these 36 intervals was checked for each of four sequences of random numbers: 20,016, 100,008, 200,016, and 1,000,008. However, in this test lag 2 (every other number) beginning with R_1 was used. $\chi_0^2 = 49.80$ for 35 degrees freedom. The resulting X^2 values are:

Sample Size	X^2
10,008	30.63
50,004	24.71
100,008	28.28
500,004	31.85

all of which are less than χ_0^2.

Test 4 - This test is the same as Test 3, and the generated sequence lengths are the same. This test differs in that the subsequence is tested using lag 2 beginning with R_2 instead of lag 2 beginning with R_1. $\chi_0^2 = 49.80$ for 35 degrees of freedom. The X^2 results are:

Sample Size	X^2
10,008	34.02
50,004	38.81
100,008	41.64
500,004	37.84

all of which are less than χ_0^2.

Poker Test

The first five digits of each random number were checked to see how many times each of the following combinations of the 16 hexadecimal digits occurred:

Combination	Probability of Occurrence
1. Bust (all digits different)	.499877929
2. One pair	.416564940
3. Two pair	.048065186
4. Three of a kind	.032043457
5. Full house	.002288818

Combination	Probability of Occurrence
6. Four of a kind	.001144409
7. Five of a kind	.000015259

Using the χ^2 test and five degrees of freedom (the last two groups were combined to give an expected frequency of at least ten), the 95 percent critical value is $\chi_0^2 = 11.04$. The resulting X^2 values are:

Sample Size	X^2
10,000	7.11
50,000	3.39
100,000	3.17
500,000	7.12

all of which are less than χ_0^2.

Serial Tests

In the following serial tests the interval 0 to 1 is divided into I equal intervals. A contingency table is constructed showing the occurrence of every interval followed by every other interval. Four kinds of hypothesis tests [$\chi^2(1)$, $\chi^2(2)$, $\chi^2(3)$, and $\chi^2(4)$] as defined previously, are performed in each of the four serial tests. For each serial test (1 through 4), there are two tables of answers, as previously defined.

Test 1: I = 10 - Generated sequence lengths = 10,000, 50,000, 100,000, and 500,000. The 95 percent critical values are:

$\chi_0^2(1)$ = 103.01 for 81 degrees of freedom

$\chi_0^2(2,3)$ = 113.14 for 90 degrees of freedom

$\chi_0^2(4)$ = 123.22 for 99 degrees of freedom.

The X^2 results are:

Test 1a:

Sample Size	$X^2(1)$	$X^2(2)$	$X^2(3)$	$X^2(4)$
5,000	74.89	85.92	86.72	98.52
25,000	79.64	94.94	83.46	98.89
50,000	54.46	68.33	60.65	74.51
250,000	60.41	69.62	77.53	86.74

Test 1b:

Sample Size	$X^2(1)$	$X^2(2)$	$X^2(3)$	$X^2(4)$
5,000	72.85	84.83	83.82	96.08
25,000	53.39	57.18	68.54	72.39
50,000	57.01	63.52	70.52	77.06
250,000	85.37	102.49	94.59	111.62

All values for Test 1a and 1b are less than the corresponding χ_0^2 values.

Test 2: I = 36 - Generated sequence lengths = 10,368, 51,840, 103,680, and 497,664. The 95 percent critical values are:

$\chi_0^2(1)$ = 1307.54 for 1225 degrees of freedom

$\chi_0^2(2,3)$ = 1343.69 for 1260 degrees of freedom

$\chi_0^2(4)$ = 1379.8 for 1295 degrees of freedom.

The X^2 results are:

Test 2a:

Sample Size	$X^2(1)$	$X^2(2)$	$X^2(3)$	$X^2(4)$
5,184	1222.4	1260.9	1274.4	1313.0
25,920	1216.8	1262.0	1246.3	1292.3
51,840	1185.6	1222.2	1212.3	1248.6
248,832	1159.0	1198.0	1185.4	1224.4

Test 2b:

Sample Size	$X^2(1)$	$X^2(2)$	$X^2(3)$	$X^2(4)$
5,184	1221.0	1266.0	1257.2	1301.0
25,920	1185.2	1215.2	1229.6	1259.8
51,840	1129.2	1156.8	1164.7	1192.4
248,832	1195.2	1221.5	1233.4	1259.7

All values for test 2a and 2b are less than the corresponding χ_0^2 values.

Test 3: I = 36 - Generated sequence lengths = 20,736, 103,680, 207,360, and 995,328. Lag 2 (every other number) beginning with R_1 is used. The 95 percent critical values are the same as in Test 2. The X^2 results are:

Test 3a:

Sample Size	$X^2(1)$	$X^2(2)$	$X^2(3)$	$X^2(4)$
5,184	1195.4	1226.3	1225.6	1258.0
25,920	1200.0	1226.7	1223.8	1250.8
51,840	1159.1	1189.0	1177.9	1207.7
248,832	1234.0	1266.6	1255.8	1288.5

Test 3b:

Sample Size	$X^2(1)$	$X^2(2)$	$X^2(3)$	$X^2(4)$
5,184	1235.2	1266.0	1267.8	1298.5
25,920	1170.7	1193.7	1197.8	1221.0
51,840	1114.3	1133.6	1145.1	1164.6
248,832	1154.1	1176.6	1186.2	1208.7

All values for test 3a and 3b are less than the corresponding χ_0^2 values.

Test 4 - This test is the same as Test 3, except that the first random number used is R_2. The 95 percent critical values are the same as in Test 2. The X^2 results are:

Test 4a:

Sample Size	$X^2(1)$	$X^2(2)$	$X^2(3)$	$X^2(4)$
5,184	1200.8	1221.2	1232.5	1252.0
25,920	1197.2	1230.8	1233.7	1267.2
51,840	1314.4	1353.4	1349.3	1388.3
248,832	1257.8	1280.4	1298.6	1321.4

The underlined values are larger than χ_0^2.

Test 4b:

Sample Size	$X^2(1)$	$X^2(2)$	$X^2(3)$	$X^2(4)$
5,184	1248.8	1276.9	1267.5	1294.0
25,920	1251.5	1286.9	1283.8	1319.3
51,840	1213.9	1247.1	1251.0	1284.2
248,832	1206.3	1248.2	1228.4	1270.3

All values for test 4b are less than the corresponding χ_0^2 values.

Runs Above and Below the Mean Test

Stated below are the results of runs above and below the mean. This test was performed on four sequences of random numbers: 10,000, 50,000, 100,000, and 500,000 random numbers.

Run Length	10,000 Random Numbers			50,000 Random Numbers		
	Above	Below	Expected	Above	Below	Expected
1	1,214	1,186	1,250	6,157	6,151	6,250
2	623	646	625	3,163	3,128	3,125
3	305	312	312.5	1,591	1,565	1,563
4	169	159	156	777	783	781
5	69	85	78	365	416	391
6	36	37	39	189	195	195
7	24	23	19.5	97	101	98
8	15	10	10	49	50	49
9	5	3	5	21	26	24
≥ 10	4	4	5	29	24	24
Totals	2,464	2,465	2,500	12,438	12,439	12,500
	100,000 Random Numbers			500,000 Random Numbers		
1	12,431	12,545	12,500	62,677	62,630	62,500
2	6,365	6,249	2,250	31,296	31,229	31,250
3	3,190	3,128	3,125	15,777	15,802	15,625
4	1,542	1,536	1,562	7,765	7,817	7,813
5	729	815	781	3,857	3,833	3,906
6	397	376	391	1,916	1,908	1,953
7	181	190	195	918	956	977
8	96	100	98	465	524	488
9	42	54	49	250	263	244
≥ 10	61	43	49	272	233	244
Totals	25,034	25,036	25,000	125,193	125,195	125,000

Runs Up and Down Test

The runs up and down test was performed on four sequences of random numbers: 10,000, 50,000, 100,000, and 500,000. The results are:

Run Length	10,000 Random Numbers			50,000 Random Numbers		
	Up	Down	Expected	Up	Down	Expected
1	2,029	2,032	2,083	10,396	10,384	10,416
2	906	934	916	4,560	4,575	4,583
3	287	254	264	1,343	1,353	1,319
4	57	72	57	284	276	288
≥ 5	18	5	12	63	57	60
Totals	3,297	3,297	3,332	16,646	16,645	16,666
	100,000 Random Numbers			500,000 Random Numbers		
1	20,732	20,692	20,833	104,073	103,804	104,166
2	9,208	9,235	9,166	45,669	45,944	45,833
3	2,665	2,695	2,639	13,176	13,194	13,194
4	546	546	575	2,935	2,908	2,877
≥ 5	127	109	119	618	621	595
Totals	33,278	33,277	33,332	166,471	166,471	166,665

Serial Correlation Test

The serial correlation coefficient r and the normal deviate t, which are functions of the random variable

$$Q = \sum_{j=1}^{k} N_j N_{j+1} ,$$

were computed for each of four sequences of random numbers: 10,000, 50,000, 100,000, and 500,000 random numbers. The 5 percent critical values are $t_0 = -1.96$ and $t_1 = 1.96$. The results are as follows:

Sample Size	r	t	u (mean)
10,000	0.007	0.766	0.49895
50,000	-0.001	-0.262	0.49939
100,000	-0.004	-1.190	0.50060
500,000	0.002	-1.480	0.50016

For all of the above sequences of random numbers, r is approximately zero and t is between -1.96 and 1.96. Therefore, the hypothesis of zero correlation is accepted and the random numbers satisfy this test.

In addition to the serial correlation results, the actual mean of each of the four sets of random numbers has been given above. As can be seen, the actual means are quite close to the expected mean of 0.5.

CONCLUSION

In this report the results of 164 statistical tests are presented. If all hypotheses were true, the probability of rejection in any single test would be .05. If the tests were independent, one could say that .05 (164) = 8.2 would be the expected number of rejections for the test series. Actually, six rejections were observed.

However, a simple comparison of six to 8.2 is not possible for the test series of this report, because individual tests are not independent here. There are two primary reasons for this lack of independence.

First, some of the tests are related in the sense that they check for the presence of the same or of similar types of nonrandom behavior. For example the four serial tests $[\chi^2(1), \chi^2(2), \chi^2(3)$, and $\chi^2(4)]$ and the serial correlation test all check for the existence of a relationship between the generated numbers, their predecessors, and their successors. If a sample fails one of these tests, it may tend to fail one or all of the others as well. This appears to be precisely what has occurred in the sample of size 51,840 in Test 4a.

The second major factor contributing to dependency in the overall test series is that the samples taken from the generated number sequence overlap in many cases. For example, if the sample sizes given for a particular hypothesis test and a particular sampling procedure are 10,000, 50,000, 100,000, and 500,000, then the sample of size 10,000 comprises the first 10,000 members of the sample of size 50,000, and so on. Thus the four corresponding tests are not independent, since the samples are not.

Other factors affecting the independence of the various tests exist, but it is felt that the foregoing two are the most significant.

Some of the dependency may be eliminated by restricting attention separately to each of the 41 (≈ 164/4) groups determined by any differences in the hypotheses tested or in the sampling procedures used. For example, the hypotheses tested in the hexadecimal digit frequency test and in the serial correlation test obviously differ, but so do those in the serial tests $\chi^2(2)$ and $\chi^2(3)$ (although the distinction is not so clear-cut in the latter instance). Also, the sampling procedures differ for frequency tests 2, 3, and 4, although the underlying hypothesis is the same in each case. Results of tests which differ in either of these two respects will be considered separately. Thus, the 41 "groups" of tests under consideration can be conveniently identified as the tests corresponding to the 41 columns of statistics in the various tables given in the Test Results section of this report. The only source of dependence present in each such group is due to the cumulative overlap of the four samples.

Assuming momentarily that every group consists of four completely independent tests of the same (true) hypothesis, all conducted at the 95 percent level, the probabilities $P(0), \ldots, P(4)$ of $0, \ldots, 4$ rejections in any one group are found (using the binomial distribution) to be $P(0) = .815$, $P(1) = .171$, $P(2) = .014$, $P(3) = .000$, and $P(4) = .000$. These numbers show (for the case of four independent tests) that the occurrence of one rejection is not unlikely, although the most probable outcome is no rejections if, in fact, the hypothesis is true.

The calculation of similar probabilities in the dependent case would be an extremely complex procedure, if a possible one. (The authors' several attempts have proven unsuccessful.) However, it appears that the real probabilities here would be only slight alterations of those given above.

The expected effect of the dependence would be to decrease the probability of one rejection and increase the probabilities of two, three, and four. Thus, some evidence

that the effect is small is that no group had more than one rejection. In particular, in the group appearing in the fourth digit position under the heading Frequency Test 1 a rejection occurred in the sample of size 50,000. Now, the dependence should be strongest between the 50,000 and 100,000 sample sizes in each group, since the highest percentage overlap (50%) occurs between these two. But in this example, not only did no rejection occur in the third test, but the calculated X^2 value for this test (12.26) is one of the lowest on the page.

Therefore, on the basis of the above probabilities and the belief that the influence of the dependence (on these probabilities) is slight, it is felt that the single rejections which occurred are insufficient evidence to reject the hypothesis tested by any group. Thus, the following conclusion seems justified: Pseudo-random numbers generated by the multiplicative congruential method, on the NAREC, possess most of the useful and characteristic properties of true random numbers and can be utilized in the ordinary applications requiring random numbers.

Although all of the numbers tested in this report were based on the first two million integers of the sequence with $x_0 = 1$ (see p. 1), there is no indication that data from any other portion of any other sequence would have produced dissimilar results for the same series of tests. However, the possibility of this remains, as well as the possibility that all or some of the numbers possess some decidedly nonrandom behavior which is undetectable by the series of tests used here. Thus, in the event that any irregularity (or perhaps we should say regularity) is observed in the subsequent use of these numbers, the authors would be appreciative if such behavior were brought to their attention.*

For copies of several articles written by other authors on random numbers or the Monte Carlo method, and for more detailed information on the results of the tests used in this report, please contact one of the authors.

ACKNOWLEDGMENTS

The authors are indebted to Frank Polkinghorn, Radio Division, to Arthur Pieper, Radiation Division, and to Dr. Herbert A. Hauptman, Dr. Benjamin Lepson, Emanuel Vegh, and many other associates of the Applied Mathematics Staff.

*Doris J. Ellis, Code 4552, NRL (Ext. 2341); Peter C. Ryan, Code 4530, NRL (Ext. 586).

REFERENCES

1. Hammersley, J.M., Handscomb, D.C., "Monte Carlo Methods," New York:Wiley, 1964, pp. 27-29

2. Kendall, M.G., and Stuart, A., "The Advanced Theory of Statistics," Volumes I and II, London:Griffin and Co., 1958 and 1961

   a. Vol. I, pp. 355-356
   b. Vol. II, p. 425
   c. Vol. II, p. 422

3. Crow, E.L., Davis, F.A., and Maxfield, M.W., "Statistics Manual," U.S. Naval Ordnance Test Station, Navord Report 3369, 1955, p. 97

4. Hoel, P.G., "Introduction to Mathematical Statistics," 2nd edition, New York:Wiley, 1954

   a. p. 173
   b. pp. 299-303

5. Kendall, M.G., "The Advanced Theory of Statistics," Volume II, 2nd edition, London: Griffin, chapter 21, pp. 125-126, 1947-1948

SELECTED BIBLIOGRAPHY

1. Brown, Bernice, "Some Tests of the Randomness of a Million Digits," The Rand Corporation, Oct. 1948

2. Good, I.J., "The Serial Test for Sampling Numbers and Other Tests for Randomness," Proc. Cambridge Phil. Soc. 49:276-284 (1953)

3. Pearson, E.S., and Hartley, H.O., editors, "Biometrika Tables for Statisticians," Volume I, Cambridge:University Press, 1954

4. Juncosa, M.L., "Random Number Generation on the BRL High-Speed Computing Machines," Ballistic Research Laboratories Report 855, Aberdeen Proving Ground, Maryland, 1953

5. Green, B.F., Jr., Smith, J.E.K., and Klem, L., "Empirical Tests of an Additive Random Number Generator," J. ACM 6:527-537 (Oct. 1959)

6. Mode, E.B., "Elements of Statistics," 2nd edition, New York:Prentice-Hall, 1951

7. Mood, A.M., "Introduction to the Theory of Statistics," New York:McGraw-Hill, 1950

8. Rand Corporation, "A Million Random Digits," Glencoe, Illinois:The Free Press, 1955, pp. xi-xxv

9. Meyer, H.A., Gephart, L.S., and Rasmussen, N.L., "On the Generation and Testing of Random Digits," WADC Technical Report 54-55, Wright Patterson Air Force Base, Ohio, 1954

10. Taussky, Olga, and Todd, J., "Generation and Testing of Pseudo-random Numbers," pp. 15-28 in "Symposium on Monte Carlo Methods," University of Florida, Mar. 1954, H.A. Meyer, editor, New York:Wiley, 1956

* * *

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*

U.S. Naval Research Laboratory
Washington, D.C. 20390

2a. REPORT SECURITY CLASSIFICATION

UNCLASSIFIED

2b. GROUP

3. REPORT TITLE

TESTS ON THE MULTIPLICATIVE CONGRUENCE METHOD OF GENERATING PSEUDO-RANDOM NUMBERS ON THE NAREC COMPUTER

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Final report on one phase of the problem.

5. AUTHOR(S) *(Last name, first name, initial)*

Ellis, Doris J., and Ryan, Peter C.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
April 2, 1965	22	5

8a. CONTRACT OR GRANT NO.

NRL Problem F04-05

b. PROJECT NO.

RR 009-03-45-5802

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

NRL Report 6217

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. AVAILABILITY/LIMITATION NOTICES

Unlimited availability
Available at CFSTI

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

Department of the Navy (Office of Naval Research)

13. ABSTRACT

This report describes the statistical tests and their results that were made on pseudo-random numbers generated by the "multiplicative congruential method" on the NAREC computer. The first two million numbers of a particular generated sequence were tested using the frequency test, poker test, serial test, runs above and below the mean test, runs up and down test, and serial correlation test. The purpose of making such tests was to determine if the generated numbers were, in fact, random in their behavior. It was found that these numbers can be considered random for most common applications.

DD FORM 1 JAN 64 1473

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Random numbers Pseudo-random numbers Computer generation NAREC Multiplicative congruential method Statistical tests for randomness						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2*a*. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through

_____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through

_____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through

_____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.